

Research Article

Leveraging spatial scale and temporal variation to track the spread rate of invasive spotted lanternflies (*Lycorma delicatula*)

Joseph Keller¹, Sebastiano De Bona², Matthew R. Helmus¹

¹ Integrative Ecology Lab, Center for Biodiversity, Department of Biology, Temple University, 1925 N. 12th Street, Philadelphia, PA 19122, USA

² Department of Biology, Bryn Mawr College, Bryn Mawr, Pennsylvania, USA

Corresponding author: Joseph Keller (jakeller@temple.edu)

Abstract

Estimating the extent and speed of invasive species' spread is crucial for optimising surveys and management, but this estimation is challenging due to the scale-dependent nature of spread. For instance, when discretising occurrence data to grids to estimate the extent of invaded ranges, larger cells can decrease boundary precision, while smaller cells can increase the risk of missing invaded areas. Moreover, on shorter timescales, such as year-to-year comparisons, spread rates can exhibit lags, accelerations and slowdowns. We present a multiscale spread optimisation methodology developed for the spotted lanternfly (*Lycorma delicatula*), a pest impacting grapes that is spreading in the USA. We analysed a dataset of > 900,000 occurrence records detailing spread from this pest's initial detection in eastern Pennsylvania in 2014 to its invasion front in Chicago, IL, in 2023. First, we delineated invaded area using square grids with varying cell sizes and α -convex hulls with different boundary resolutions. For each method and scale, we regressed invaded range area against year using both simple linear and logistic non-linear models. Coarser spatial scales yielded faster estimated spread rates, and logistic models outperformed linear regressions, indicating a lag, acceleration and slowdown in annual spread rate. Next, to determine the spatial scale that best captured the invasion front boundary, we used cross-validation, optimising the F_β metric, which balances recall and precision. Emphasising recall generally favoured coarser spatial scales and α -convex hulls outperformed grid-based methods. Finally, *L. delicatula* has dispersed long distances from its primary epicentre, establishing satellite populations. We delineated each satellite population each year since its inception using optimised α -hull values and measured their year-to-year boundary expansion. Like the primary epicentre, these satellite populations showed accelerating expansion, at least until merging with the primary invasion. Overall, the expansion rate of the primary invasion has slowed since 2017, decreasing from a maximum of 21 km/year, possibly due in part to control efforts, although the exact cause remains unresolved. Considering spatial scale and temporal variation can improve the analysis of invasive spread and motivate early interventions to manage nascent epicentres before their spread accelerates.

Key words: Alpha hull boundary, grape pest, integrated pest management, invasion curve, invasion front delineation, invasive species, jump dispersal, range limit size, spatial scale, spread rate



Academic editor: Samuel Ward

Received: 4 November 2025

Accepted: 10 March 2026

Published: 27 March 2026

Citation: Keller J, De Bona S, Helmus MR (2026) Leveraging spatial scale and temporal variation to track the spread rate of invasive spotted lanternflies (*Lycorma delicatula*). NeoBiota 106: 287–307. <https://doi.org/10.3897/neobiota.106.177041>

Copyright: © Joseph Keller et al.

This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

The spread patterns of invasive species strongly influence their economic impact (Epanchin-Niell and Liebhold 2015). Detailed knowledge of how far and fast the targeted outbreak is spreading can guide the choice of the most appropriate

management objectives, which can range from eradication, to slowing spread, to limiting impacts in invaded areas (Rejmanek and Pitcairn 2002; Liebhold and Kean 2019). Additionally, accurate knowledge of spread rates can improve the specific strategies deployed to achieve those objectives, guiding optimal spatial and temporal allocation of surveillance and management efforts (Andrew and Ustin 2010; Epanchin-Niell and Hastings 2010; Triska and Renton 2018).

Temporal variation complicates the process of characterising the spread patterns of invasive species. Theory predicts that invasions' spread rates should generally follow a predictable pattern over time (Shigesada et al. 1995). Initially, spread is expected to be slow, as populations take time to grow; lag phases between initial establishment and subsequent rapid spread are, in fact, frequently observed (Crooks and Soulé 1999; Memmott et al. 2005; Coutts et al. 2018). After this lag phase, spread accelerates and may approach a constant asymptotic rate for some time, depending on the specific characteristics of the invasion (Hastings et al. 2004). Finally, the spread rate in the invaded range must slow as more and more of the potentially suitable range becomes occupied (e.g. Ward et al. (2020)). As the invaded range becomes saturated, dispersing propagules are more likely to end up in areas that are already occupied or unsuitable for survival, resulting in no further spread. The emerald ash borer (*Agilus planipennis*), for example, has had varying spread rates in the USA over time, with breakpoint analysis revealing four distinct spread periods that followed the expected slow-rapid-slow pattern closely (Ward et al. 2020). The spread rate of the spongy moth (*Lymantria dispar*) in North America has also varied over time, though less predictably, with rapid spread between 1900 and 1915, followed by a slower spread phase from 1916 to 1965 and another rapid spread phase between 1966 and 1990 (Liebhold et al. 1992). Managers benefit from a detailed understanding of spread rates' variation and from information on the status of the targeted invasion.

The need to select an appropriate spatial scale for analysis further complicates efforts to characterise invasive species' spread patterns. Methods for characterising these patterns rely on delineating invaded areas from uninvaded areas, a process inherently dependent on scale. Estimation of the size of the invaded range and the location of the invasion front depend on the spatial scale used in the analysis. For example, invasion fronts are drawn based on geopolitical units (Tobin et al. 2007), trap locations (Tobin et al. 2013), point locations of reported presences (Wallace et al. 2020), remotely sensed data (Lonsdale 1993) and expert opinion (Storer 1937). Regardless of the approach, at coarser scales, species can be said to occupy entire nations, while, at much finer scales, the spatial footprint of each individual will make up the species' distribution. Scaling must be considered whether the invaded range is evaluated based on areal measures (grid cells or geopolitical units occupied) or boundary drawing (where the accepted level of concavity and the inclusion/exclusion of small outlying populations and/or unoccupied pockets depend on scaling parameters). With the growing availability of geolocated point records describing the occurrence of organisms, identifying an appropriate scale for analysing such data is becoming increasingly important (Shirey et al. 2021).

Given these issues of spatial scale and temporal variation, researchers have used various methods to measure spread patterns. For example, a common approach estimates spread rate by tracking changes in the invaded area over time, approximating the area as a circle and calculating annual changes in its radius using square-root area regression (Skellam 1951; Hastings et al. 2004). Alternatively, spread rates may be

estimated by measuring the distances between invasion fronts from consecutive time steps. Vectors between boundaries may be measured radiating from a central point (typically the origin of the invasion) (Tobin et al. 2007), orientated perpendicular to the invasion front or drawn as the shortest distance between points (Tisseuil et al. 2016). In general, without a clear consensus on which method is best to estimate spread patterns, authors often apply multiple approaches to the same spatio-temporal datasets (e.g. Liang et al. (2019); Ward et al. (2020); Cook et al. (2021)).

Well-documented invasions provide valuable case studies to test and compare methods for estimating spread patterns (Hastings 1996). The spotted lanternfly (*Lycorma delicatula*) is an invasive planthopper whose spread in the USA has been thoroughly surveyed through the combined efforts of state and federal agencies (De Bona et al. 2023). A univoltine phloem feeder, the spotted lanternfly damages vineyards by feeding on grapevines and has spread rapidly in the USA since its first detection in Berks County in south-eastern Pennsylvania in 2014 (Cook et al. 2021; Urban and Leach 2023). Satellite populations are common because this pest hitchhikes on a variety of transported materials and readily establishes (Huron et al. 2022). This pest feeds on a wide variety of plants (Barringer and Ciafré 2020), but has a preference for the invasive tree *Ailanthus altissima* (Urban and Leach 2023), which is also widespread in the United States. Here, we present a detailed analysis of the spread patterns *L. delicatula* has exhibited thus far in the United States, focusing on the following questions:

1. How has the spread rate varied over time?
2. How does spatial scale affect spread rate estimates?
3. How do different methodologies for delineating invaded ranges affect spread rate estimates?
4. What spatial-scale parameters result in best-fitting delineations, based on cross validation?
5. Do satellite populations follow similar patterns of spread as the primary invasion?

Materials and methods

We evaluated change in the area of the invaded range using both grids and α -convex hulls and assessed evidence for temporal variation in spread rates by comparing the fit of simple linear regressions and logistic curves. We identified best-performing spatial scales for analysis using cross validation and, based on those parameters, measured spread rates around satellite population epicentres that have arisen over time. These results provide an updated understanding of *L. delicatula*'s spread in North America.

Data

Exceptionally high effort has been devoted to locating invasive spotted lanternfly populations in the United States. Trained professionals have deployed traps and conducted widespread visual surveys. Public awareness campaigns highlighting this relatively large and visually distinctive pest have yielded many reported sightings made by members of the public. To characterise the growing distribution of *L. delicatula* in the United States, we gathered presence and absence records documenting the sampling effort from state agencies and the USDA covering the period from 2014, when the pest was first found, through 2023 (Table 1). These

Table 1. Sampling effort has varied over time as the *L. delicatula* invasion in the USA has continued to grow. Numbers of presence and absence records for *L. delicatula* from government (state and federal) sources over the biological years 2014 through 2023 are listed, along with two measures of the invasion size: the invaded area and the length of the boundary, both as assessed with an α -convex hull with $\alpha' = 35$ km. Biological years turn over on 1 May. To show how sampling effort compared to the size of the invasion, we report the number of observations per km² area and per km of boundary perimeter.

Biological year	Number of presences	Number of presences with confirmed establishment	Number of absences	Invaded area (km ²)	Invasion boundary length (km)	Sampling effort (Obs. per km ² area)	Sampling effort (Obs. per km boundary)
2014	62	26	308	33.5	26.3	10.6	14.1
2015	2,876	2,479	4,800	274.0	77.3	28.0	99.4
2016	2,780	2,286	6,482	1718.1	188.3	5.4	49.2
2017	1,333	895	7,843	7200.7	403.5	1.3	22.7
2018	4,705	3,937	95,552	16387.6	644.1	6.1	155.7
2019	10,441	7,477	120,535	29708.3	1080.5	4.4	121.2
2020	10,867	3,765	65,985	39980.6	1688.2	1.9	45.5
2021	10,799	5,116	53,944	54924.1	2535.9	1.2	25.5
2022	12,593	6,812	52,853	83808.6	3487.0	0.8	18.8
2023	8148	5,951	32,803	112695.3	3989.9	0.4	10.3

records were included in anonymised form in the lydemapr database (De Bona et al. 2023). Absences in this database reflect negative search or trapping records. All analyses were based on biological years reflecting the annual life cycle of *L. delicatula*. Each biological year’s records reflect the presence or absence of individuals that hatched in that year and the eggs they produced. Egg masses laid in one calendar year but discovered before hatch in the next, for example, would be included in the first year’s data. In practice, the biological year turns over on 1 May. Data were anonymised as in De Bona et al. (2023). We based our analyses on records of established *L. delicatula* (indicated by two or more living *L. delicatula* individuals or an egg mass) in order to exclude regulatory incidents where either single individuals or dead individuals were correctly reported as present, but did not reflect the presence of a surviving and/or reproducing population of *L. delicatula*.

Grid-based boundaries

To distinguish invaded areas from uninvaded areas, we divided the eastern USA into grids of varying scales and assessed the presence of *L. delicatula* within each cell. In each year, cells where at least one established *L. delicatula* population was noted in that year or in any prior year were considered invaded, while cells without any of these records were considered uninvaded. Records of established *L. delicatula* were represented as points and included results from visual observations and tree-based traps. While we used regular grids with varying resolution here, this process is closely analogous to methods based on geopolitical units such as townships, counties, states and countries (Liang et al. 2020).

α -convex hull boundaries

To test an alternative method for delineating invaded and uninvaded areas, we fit α -convex hulls. α -convex hulls are mathematically well-defined (Edelsbrunner et al. 1983) and are capable of accurately locating range boundaries for irregularly-shaped and disjointed ranges (Burgman and Fox 2003). This approach to locat-

ing species range boundaries has been applied most frequently in investigations of the range size of threatened species (IUCN Standards and Petitions Committee 2024), though it has also been used previously to estimate range sizes of potential invasives (Hui et al. 2011) and to track the spread of biological invasions (Urban et al. 2008). α -convex hulls are constructed using triangulation of presence points, with a single parameter (α) determining the resolution of the boundary, which consists of arcs of open circles with a radius equal to α (Pateiro-Lopez and Rodriguez-Casal 2022). When α is very small, each presence is included as a disjointed point, and when α is infinitely large, the α -convex hull is congruent with the minimal convex hull (Edelsbrunner et al. 1983). For intermediate α values, polygons with varying resolution are generated. We used the *ahull* function from the *alpha-hull* package (Pateiro-Lopez and Rodriguez-Casal 2022) and the *edmaps* package (Camac et al. 2021) in R (R Core Team 2023) to fit and process α -convex hulls.

To aid comparison with the grid-based analysis, we evaluated α values that resulted in circles with area equal to that of the square grid cells we evaluated. To do this, we report scaling results using a derived parameter α' , which is simply the edge length of a square with equal area to the circle with radius α , calculated as follows:

$$\text{grid cell edge length} = \alpha' = \sqrt[2]{\pi} * \alpha \quad (1)$$

For our analyses, we evaluated spread patterns produced using cells with grid edge lengths and α' values of 1, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 75 and 100 km (19 in total). Pennsylvania counties have a mean area of 1780 km², which corresponds to an edge length of 42.2 km. Our values, therefore, range from much finer than the county scale to much coarser (though county sizes in other parts of the country vary). We additionally included $\alpha' = 1000$ km to approximate the minimum convex hull. In a few cases, hull fitting failed due to linearity of points in small clusters. In these cases, we jittered points randomly up to 10 m away in both northing and easting and refitted the hull.

Square-root area-time regression

We employed square-root area-time regression to estimate *L. delicatula* spread rate through time in the USA. We used the *st_area* function from the *sf* package in R to measure the area of the invaded range as determined by each boundary drawing method and parameter setting (Pebesma et al. 2023). Then, we calculated the effective range radius (ERR) (Shigesada et al. 1995):

$$\text{ERR} = \sqrt{\frac{A}{\pi}} \quad (2)$$

where A is the invaded area. ERR is the distance from the centre to the edge of a circle with equivalent area to the invasion. For each bounding method and scale, we fitted simple linear models regressing the effective range radius against year. We also fitted logistic non-linear curves following the form:

$$\text{ERR} \sim \frac{\text{asymptote}}{1 + e^{(\text{mid}-t)/\text{slope}}} \quad (3)$$

where t is the year, *asymptote* represents the eventual upper limit, *mid* represents the time at which growth is the fastest and *slope* is a parameter controlling the slope of

the curve. We used the *nls* and *SSlogis* functions in the *stats* package in R (R Core Team 2023). The slope of the linear model is the estimated spread rate. For the logistic curve, spread rate varies over time and is estimated by tangents to the curve. We evaluated models' fit to the data based on AICc, the small sample corrected Akaike Information Criterion (Burnham and Anderson 2002).

Cross validation to select optimal scale for invaded range delineation

Both the grid-based and the α -convex hull methods for bounding the invaded range require the user to select a scale parameter. For grid-based bounding, the user selects the length of the cells' edge. For α -convex hulls, the user chooses parameter α , which controls the boundary's resolution. We tested 19 and 20 parameter values, respectively (listed in the previous section), with these two methods and used cross validation to identify which parameter values resulted in boundaries that most effectively distinguished locations of *L. delicatula* presence from locations of absence. First, we randomly split the dataset into five subsets. We then performed cross validation separately for grid-based bounding and α -convex hulls (*methods*), for each year from 2014 to 2023 (*focal year*) and for each of the tested parameter values (*parameter*). For each combination of method, focal year and parameter, we generated boundaries using a dataset that excluded one of the five subsets (the excluded subset later used as test data). Presence records from the year under investigation and from all prior years were used to create the boundary. Then, we evaluated how effectively the resulting boundary captured presences and excluded absences by comparing it to the withheld test data (see Suppl. material 1: fig. S1.1). We used the portion of the test data from the focal year to evaluate both the precision (proportion of points included in the boundary that were truly presences) and recall (the fraction of all presence records that were included in the boundary). To balance the number of presences and absences used in evaluation, we resampled absences to ensure equal sample sizes. We note that imperfect detection may underlie some of the recorded absences, rather than lack of colonisation. Nonetheless, improving precision remains a valuable goal, even with potential failed-detection absences present in the dataset.

We calculated recall and precision for each combination of method, focal year and parameter, summing false positives, false negatives and true positives across the five folds. Then for each method, year and scale, we calculated F_β scores:

$$F_\beta = (1 + \beta^2) * \frac{\text{precision} * \text{recall}}{(\beta^2 * \text{precision}) + \text{recall}} \quad (4)$$

where β is a positive value that reflects the emphasis placed on recall, such that recall is β times as important as precision (Chinchor 1992). The F_β score provides a flexible metric for evaluating the scale-dependent delineations of invaded area from uninvaded area that reflects the decisions made by stakeholders (Hauser et al. 2016). This metric ranges from 0 to 1, with 1 reflecting perfect precision and recall. We evaluated scales' performance using β values ranging from 1 (equal weight for presence and recall) to 4 (four times the emphasis placed on recall). Varying β reflects differing levels of importance that stakeholders may place on accurately capturing all presence points (high recall, emphasised with higher β) versus only indicating presence at specific locations where the targeted pest is found (high precision). Decision-makers may determine the relative weight given to each of these goals by selecting differing β values. We discuss alternative metrics in the

discussion, but note that, because bounding methods result in binary predictions rather than continuous values which could be split based on a threshold, we were unable to use metrics that depend on continuous scores, such as area under the receiver operating curve (Lawson et al. 2014). We calculated F_β scores for each year of observation and with contingency table values summed across years, giving an estimate of the overall preferred values of cell edge length and α' .

Irregularity in the shape of the invaded range may vary with time and with spatial scale. To better understand the influence of the invaded range's shape on the selection of a preferred scaling parameter, we calculated an index of circularity for each α -convex hull as the inverse of the normalised perimeter. Normalised perimeter is calculated by dividing a shape's perimeter by the perimeter of a circle of equal area (Patton 1975). This index of circularity ranges from 0 to 1, with low values indicating a shape differs dramatically from circular, while a value approaching 1 indicates similarity to a circle. Hulls were clipped to the coastline to remove any areas falling in the ocean prior to calculating this metric.

Satellite population spread analysis

Multiple distinct satellite populations of *L. delicatula* have arisen in the USA after long-distance dispersal events. These distinct satellite populations provide additional cases where the early years of spread after introduction can be investigated. We fit α -convex hulls to the distribution of *L. delicatula* in each year, using the overall best performing value of α' with $\beta = 2$ (35 km) and identified all distinct satellite populations by checking all polygons and seeing which overlapped each other in consecutive years. To focus on adequately sampled outlying populations, we excluded from analysis any outlying populations with fewer than 20 recorded absences within a 5 km buffer around their margins in each year. We measured each satellite population's area in each year that it was present using the *st_area* function in the *sf* package (Pebesma et al. 2023).

When tracking the spread around distinct *L. delicatula* populations rather than the overall growth of the invasion as a whole, boundary displacement measurements can offer a clearer picture of spread than area-based metrics (Tobin et al. 2007). This difference arises because small outlying populations are frequently subsumed by the large initial population. In our analysis, in cases where two or more formerly distinct populations merged, the oldest joining population was said to continue and any younger populations to no longer exist as distinct units. With this protocol, when an outlying population is enveloped by an older population, the older population appears to grow by the entire area of the outlying population (including previously occupied area, see Suppl. material 1: fig. S1.2). The size of the older population then increases substantially even if new spread was relatively little. By focusing on boundary displacement, however, we focus on actual spread to new locations.

We measured distances between consecutive years' boundaries for all epicentres. While the α -convex hull provides a useful tool for bounding points where *L. delicatula* is established, it does not result in a "regular" border as described in Sharov et al. (1995), that is, a border where all lines radiating from an interior point cross the boundary only once. Islands, folds, and gaps in the boundary remain. For each cluster, we evenly spaced 300 points around the boundary for a given focal year, t , using the *st_line_sample* function (Pebesma et al. 2023). Then, for each of these points, we measured the shortest distance from that point to the border of

any boundary polygon from the previous year, $t-1$, using the *st_distance* function. This approach accounted for instances when two or more disjunct populations merged from year to year. In these cases, the boundary displacement distance was measured to whichever of the two clusters was closer in the prior year.

Results

The full database contained 490,211 records of *L. delicatula* presence and 441,012 records of absence, including public records (Table 1). Of these points, 38,744 were records of *L. delicatula* establishment (indicated by a minimum of two live individuals or the presence of an egg mass) made by state or federal surveyors. Established *L. delicatula* populations have been reported as far north as 43.21 °N, as far south as 35.63 °N and as far west as 87.64 °W (locations with established *L. delicatula* populations are shown in Fig. 1). The primary invasion has grown to cover large portions of Pennsylvania, New Jersey, Delaware and Maryland. By 2023, far-flung outlying populations were established in several additional states, including Michigan, Illinois, Tennessee, North Carolina and Kentucky.

Square-root area-time regression

Spread rate estimates were substantially higher at coarser spatial scales (Fig. 2, see Suppl. material 1: table S1.1–S1.6). This was true for both grid- and α -convex hull-

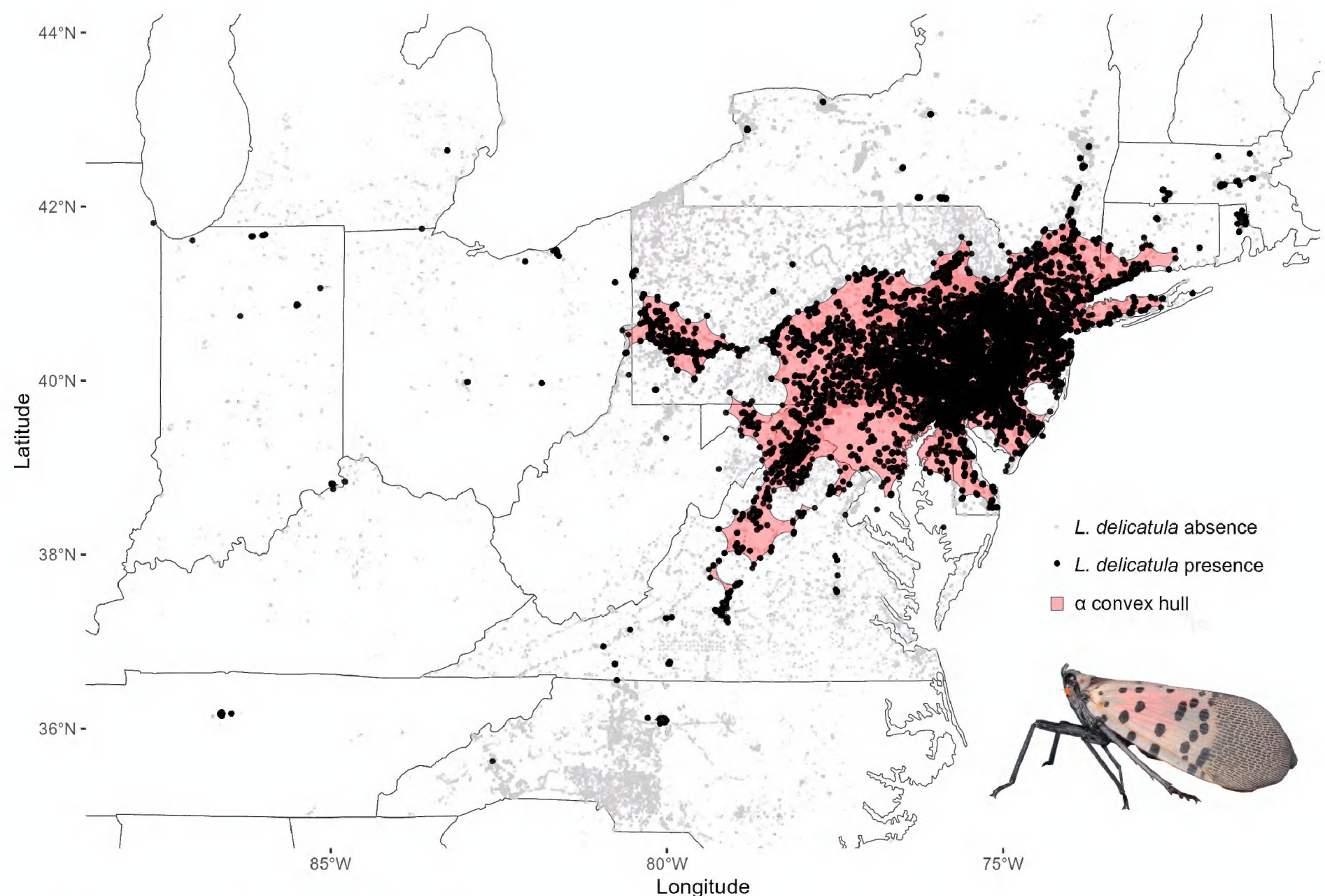


Figure 1. *Lycorma delicatula* (spotted lanternfly) occurrences in the United States as of 2023. Surveyed properties where government agencies have documented *L. delicatula* establishment are shown as black occurrence points, while those with no *L. delicatula* discovered are light grey. The invaded range (red polygon) was estimated as an α -convex hull (best-fit $\alpha' = 35$ km).

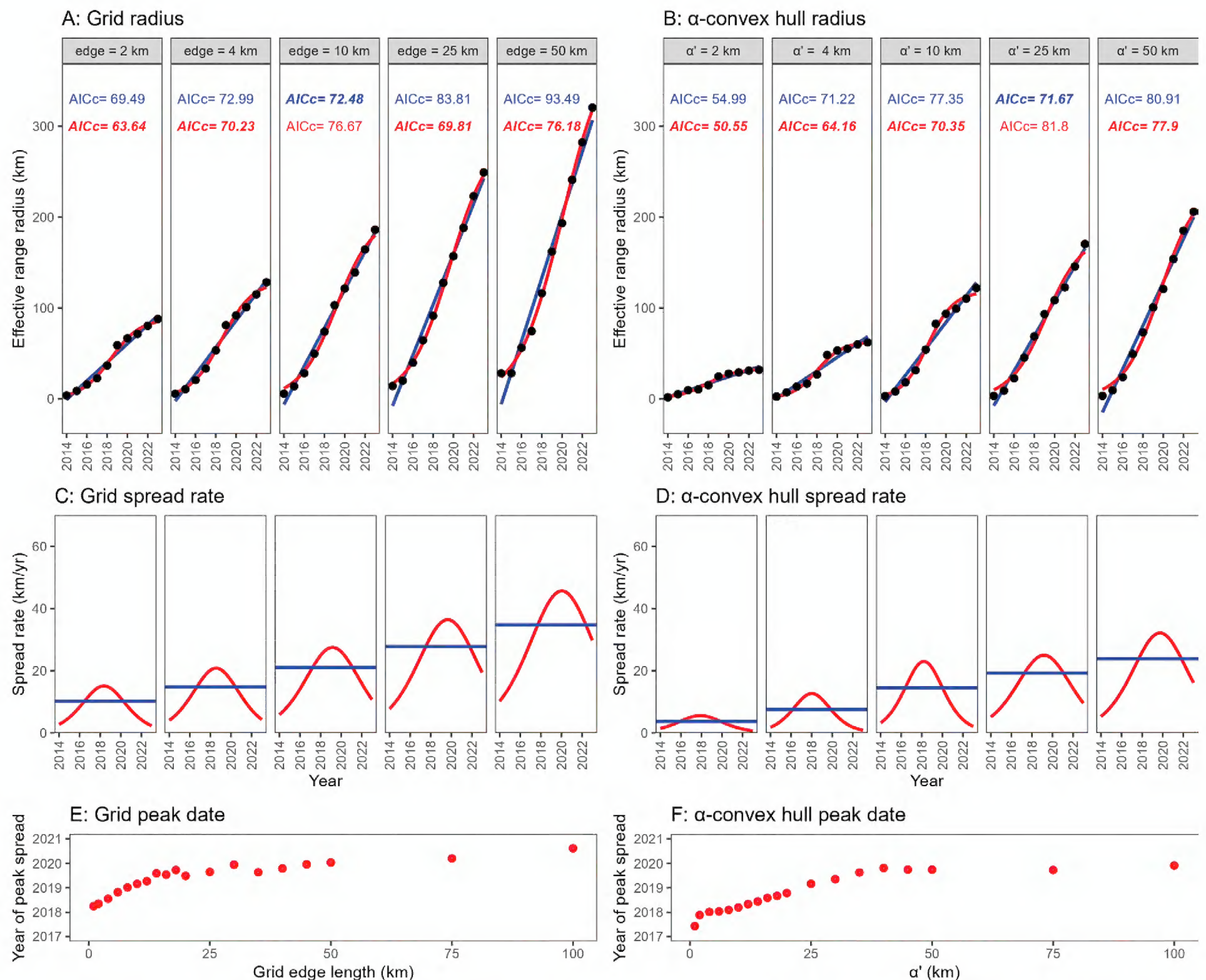


Figure 2. *Lycorma delicatula*'s estimated spread rate increased with spatial scale (grain size). Invaded range area was estimated based on grids (A, C, E), for which grid cell edge length varied and α -convex hulls (B, D, F), for which hull resolution varies (i.e. α' , eq. 1, Methods). To estimate spread rate from model slopes, we converted range area to effective range radius (radius of a circle of equal area) and fitted simple linear regression models (blue) and logistic curve models (red), regressing ERR against year. We evaluated their fit using AICc (A, B). For most scales, logistic models fit better than linear models. Coarser scales of occurrence aggregation had higher spread rates (C, D) and later timing of peak spread as estimated from the logistic models (E, F), though the date of peak spread plateaued around 2020 as spatial coarseness increased.

based methods of bounding the invaded area. With coarser scaling parameters, *L. delicatula*'s invasion in the USA was estimated to cover a larger area, resulting in larger values for the spatial extent of the invasion (ERR) and more rapid increase in ERR over time (Fig. 2). Generally, the fitted logistic curve had better fit to the data than the simple linear regression as measured by AICc (Fig. 2, Suppl. material 1: table S1.1–S1.4). The varying slope of the logistic curve precludes reporting a single spread rate, but the maximum spread rate offers a useful comparison across spatial scales. Across the range of scales tested, the maximum spread rate increased from 9.6 km/year (with edge length 1 km) to 57.6 km/year (with edge length 100 km) for grids and from 2.4 km/year (with $\alpha' = 1$ km) to 39.9 km/year (with $\alpha' = 100$ km) for α -convex hulls (Fig. 2C, D). Coarser scales also resulted in later estimated peak spread rate timing. For both grid and α -convex hulls, as scale increased, the peak shifted from occurring in 2018 to occurring in 2020.

Cross validation to select optimal method and scale for invaded range delineation

The α -convex hulls generally delineated the location of the invasion front more accurately than grid-based boundaries, as measured by F_β score based on 5-fold cross validation. With $\beta = 1$, the best-performing α -convex hull fitted the test data better than the best-performing grid-based boundaries for each year (Fig. 3A). Differences between grid performance and α -convex hull performance were larger at small values of β and the two methods' performance was comparable with $\beta = 3$ or 4. Due to the superior or comparable performance of α -convex hulls, we focus on this method for the remainder of cross validation results.

The α' value that performed best increased over time (Fig. 3B, compare x-value of shaded regions moving across rows). In early years, the invasion covered relatively little area and formed one contiguous shape, and sampling effort was dense (Fig. 2, Table 1). Under these conditions, small α' values were favoured (Fig. 3B). By 2023, however, the invaded area was discontinuous and irregularly-shaped, and sampling density was lower. Under these conditions, higher α' values resulted in better performance. Higher β values resulted in coarser preferred spatial scales (Fig. 3B, compare x-values of the shaded region within columns). For the α -convex hull analysis, the optimal scale, based on averaging across years, ranged from $\alpha' = 12$ km with $\beta = 1$ to $\alpha' = 100$ km with $\beta = 4$.

For further analyses, we selected the optimal value of $\alpha' = 35$ km, based on the pooled contingency table results across years with $\beta = 2$. Choosing $\beta = 2$ in the F_β score emphasises recall over precision, making this β value particularly useful for identifying boundaries that prioritise capturing all potential presences. Thus, a $\beta = 2$ is advantageous when it is critical to avoid underestimating the invaded range and consequently missing outlying populations.

Satellite population spread analysis

We identified 153 distinct satellite populations over the course of *L. delicatula*'s invasion in the USA, based on α -convex hulls drawn with $\alpha' = 35$ km (Fig. 4). Of these, 46 persisted as distinct, disjointed invasion epicentres for at least 1 year without being subsumed into the primary invasion. However, not all satellite populations were sampled intensively enough to characterise year-to-year boundary displacement. Specifically, only 29 of the 1-year-old epicentres were sufficiently sampled, having at least 20 recorded absences within 5 km of presences in each year of observation (years 0 and 1). Overall, the epicentres exhibited slow initial spread, and, in several cases, the invaded range did not change between the year of first discovery and the next, despite continued sampling around the selected satellite populations. The lack of observed spread, given the existing sampling effort, suggests any such spread was limited. Median boundary displacement distance for satellite populations in their first year ranged from 0 km to 10.6 km. In their second year, boundary displacement was larger, generally, with median displacement ranging from 0 km to 15.4 km.

Discussion

Analysis of the well-documented record of *Lycorma delicatula*'s distribution over time in the USA revealed that the invasion's spread rate has varied over time, generally following the slow-fast-slow pattern expected for biological invasions (Fig. 2)

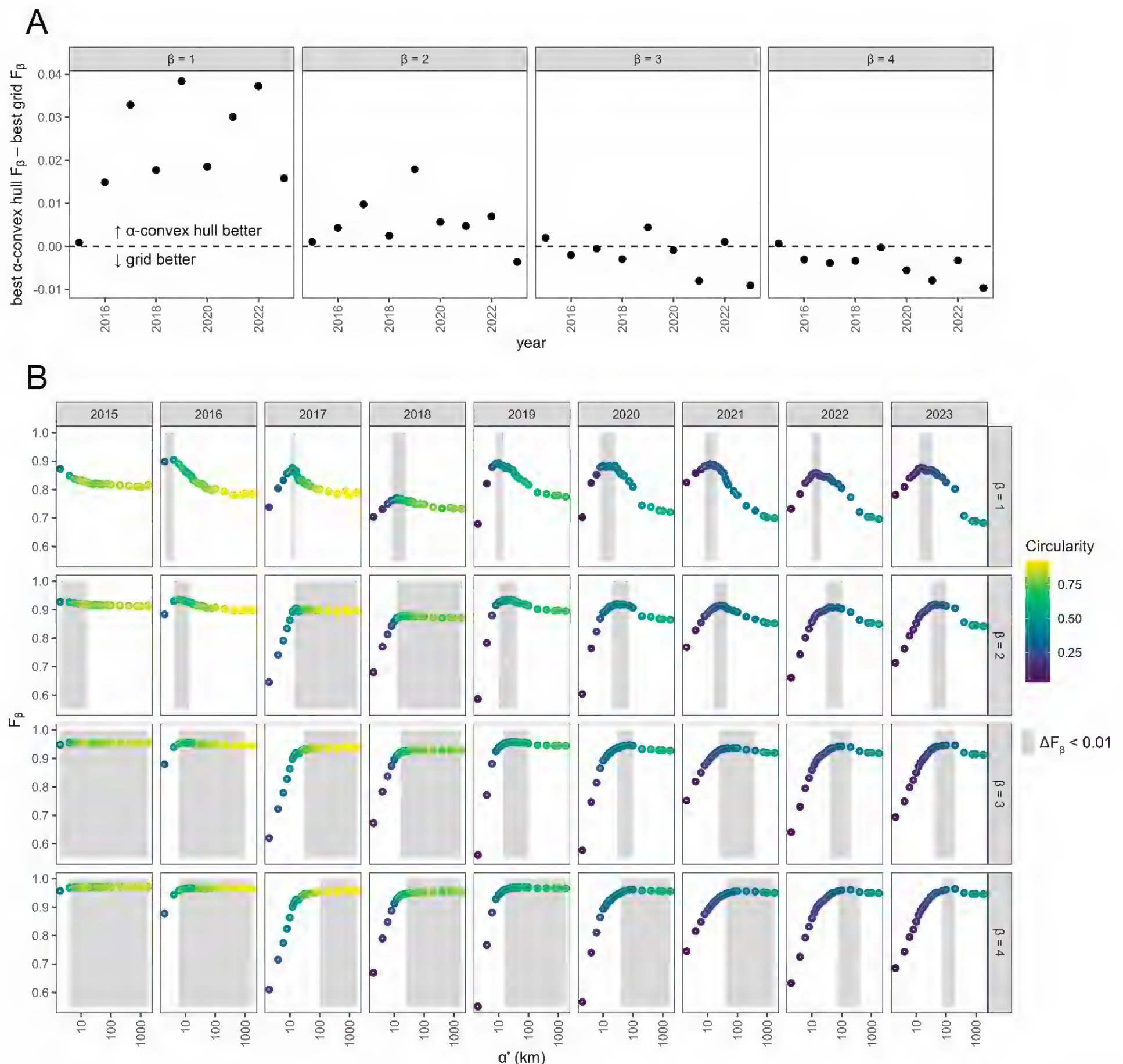


Figure 3. Best performing spatial scales varied over time and with the relative importance placed on recall versus precision. Using 5-fold cross-validation, we evaluated *L. delicatula* invasion boundaries in the USA annually, comparing grids with varying cell-edge lengths to α -convex hulls with varying α' resolution values (eq. 1). Boundaries were assessed using the F_β score, with β from 1 (equal weight on recall and precision) to 4 (quadruple weight on recall). α -convex hulls generally outperformed grids, except at high β where performance was comparable (A). The F_β scores for each year (columns, panel B) and β value (rows in panel B) were calculated, with the top-performing α' values marked as shaded areas (within 0.01 of observed maximum) Best-performing α' values increased with β (shaded areas shift right moving down columns). In early years when the invasion was more circular (yellow points), many values of α' performed similarly, indicated by the wider shaded areas in the lower left panels.

with peak yearly boundary displacement of around 21 km/yr. The > 900,000 presence and absence point records allowed us to evaluate the influence of spatial scaling and employ cross validation to select an optimal spatial scale to estimate spread patterns (Figs 1, 3). Analysis with coarser scales resulted in faster spread rate estimates (Fig. 2). Generally, finer scale parameters were favoured in early years when sampling density was high and the distribution was compact and contiguous, while intermediate scaling parameters performed better in later years when sampling

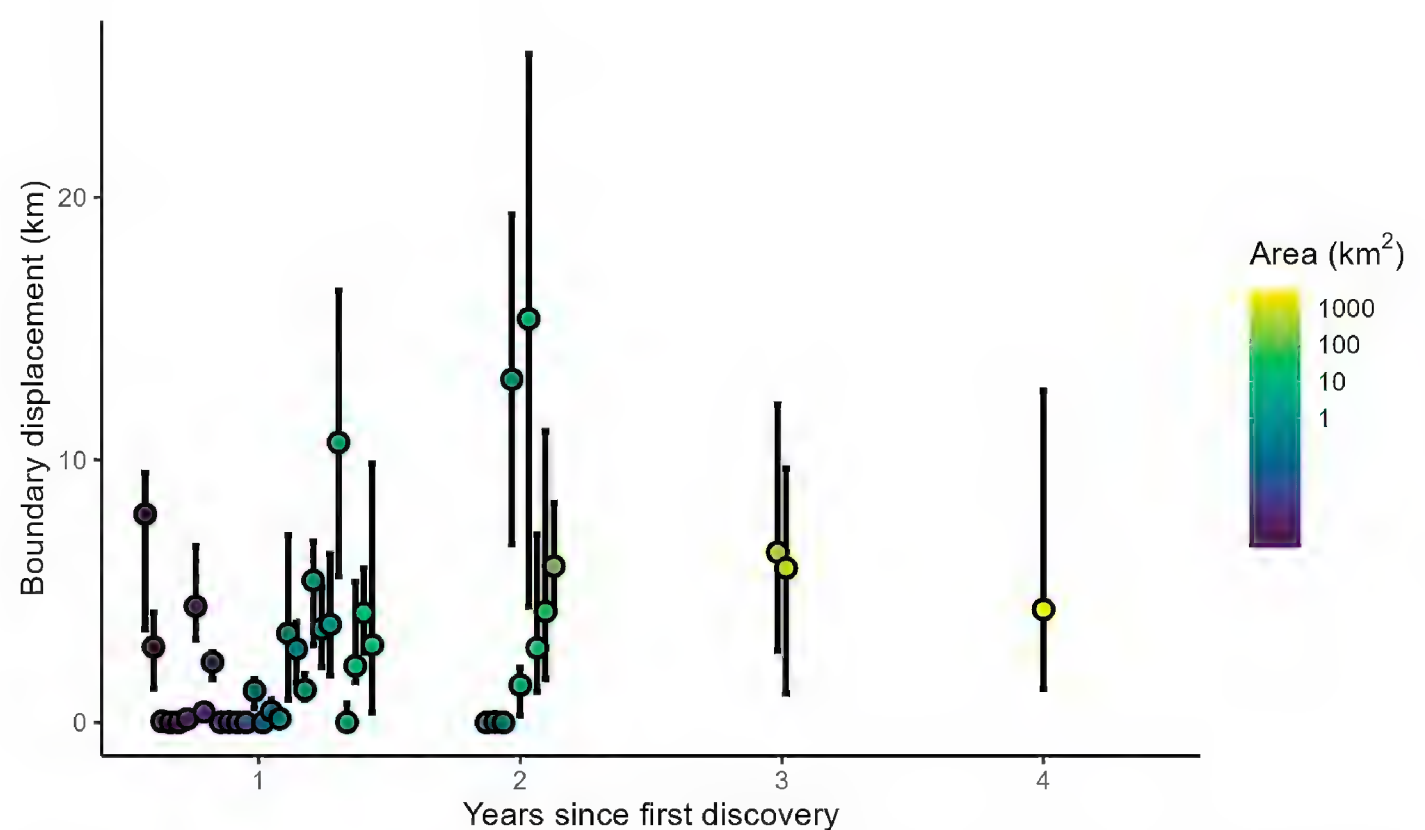


Figure 4. Spread rate varied with satellite population age. Distances between consecutive year boundaries of *L. delicatula* satellite populations were measured using α -convex hulls. Most epicentres located near the primary invasion were subsumed within a few years. Points represent median boundary displacement, with bars showing the inter-quartile range. Points are coloured, based on the population's area in the prior year. In the first year after detection, most epicentres exhibited shorter spread distances compared to the > 20 km annual spread observed for the primary invasion 3–5 years after its detection.

density was lower and the distribution discontinuous and irregular (Fig. 3). Most satellite populations had median boundary displacement under 1.5 km in the first year after discovery (Fig. 4). Together, these findings demonstrate the importance of considering temporal variation and spatial scaling when deriving spread pattern estimates and highlight the importance of targeting nascent satellite populations with management before their spread rates accelerate to the rapid rates observed around primary invasions.

We observed rapid annual spread around the primary *L. delicatula* invasion during the years of peak spread, with boundaries moving a median distance of 20.7 (IQR: 14.5 – 30.9), 20.3 (IQR: 12.0 – 31.3) and 20.5 km (IQR: 13.1 – 31.1 km) in 2017, 2018 and 2019, respectively. The spread of *L. delicatula* in the USA has been comparable to that of other introduced insect pests. The spongy moth (*Lymantria dispar*), for example, was estimated to spread at an average rate of 21 km yr⁻¹ from 1966 to 1990 before the implementation of a spread containment programme (Liebhold et al. 1992). The Japanese beetle (*Popillia japonica*) was estimated to spread at 11.9 km/year for 1939–1951 (Allsopp 1996). Finally, a median spread rate of 17.7 km/year (range 0.014 to 400 km/year) has been observed across 86 insect species introduced to the USA (Fahrner and Aukema 2018).

For *L. delicatula*, future work should focus on explaining variation in spread patterns, based on land cover, transport flows and environmental conditions (Andrew and Ustin 2010; Tobin and Robinet 2022). Prior analysis of *L. delicatula* spread rates found an association between human population density and *L. delicatula* spread, with denser human populations increasing the likelihood that counties were colonised (Cook et al. 2021). Models of *L. delicatula* spread incorporate the influence of dispersal pathways including rail, transport corridors and trade connections, which can potentially help individuals to move long distances and

colonise new areas (Jones et al. 2021; Huron et al. 2022; Ladin et al. 2023). Further, *L. delicatula* population dynamics are temperature-dependent and this effect should be explicitly incorporated into future research explaining spread patterns (Kreitman et al. 2021; Lewkiewicz et al. 2022, 2024). The spread of *L. delicatula* has involved repeated long-distance transport events (Fig. 1) (Urban et al. 2021; Belouard et al. 2025). Future research should focus on identifying environments that increase the likelihood of generating propagules capable of transport. Additionally, understanding the types of environments that support the establishment of new satellite population epicentres could enhance surveillance strategies.

Cross-validation is a robust method for identifying an appropriate spatial scale for spread pattern analyses. Historically, data availability often constrained the choice of spatial scale; for instance, data documented at the county level effectively set the minimum resolution for analysis (Liang et al. 2020). Today, however, the widespread availability of geolocated point data – provided by citizen scientists contributing to online platforms and by trained surveyors using accurate and accessible GPS technology – offers finer resolution for mapping invader distributions (Crall et al. 2011; Nathan et al. 2022). Despite these advancements, researchers must still devise methods to delineate invaded from uninvaded areas. The goal is to establish boundaries capable of accurately distinguishing additional presence and absence locations if further observations were made. Cross-validation simulates this process by holding aside a subset of the data to evaluate how well boundaries drawn from the remaining data predict the withheld records. Future work should explore how different cross-validation strategies influence the selection of spatial scales for spread pattern analyses, particularly in the context of highly autocorrelated, heterogeneous or fragmented landscapes (Dewhurst and Lutscher 2009; Meyer et al. 2019). Investigating the impact of varying the proportion of data withheld and the spatial arrangement of withheld points could provide insights into the robustness and generalisability of boundary delineation methods (Roberts et al. 2017). We note, however, that the process of splitting data into spatial folds itself requires the choice of a scale at which to group points, adding a new dimension along which to optimise and compounding the complexity of the selection process. Finally, extending cross-validation techniques to better account for temporal lags and spread rate variation could enhance predictions of invasion front progression over time, providing more reliable tools for management and surveillance (Neubert and Parker 2004; Aikio et al. 2010; Lampert 2024).

Comparing the performance of different spatial scales when delineating invaded ranges requires selecting an evaluation metric, with the F_β score being particularly useful because it encapsulates two key properties relevant to invasion biology: precision and recall. Precision is the proportion of true positives (correctly predicted presences) out of all predicted presences (true positives plus false positives). High precision minimises false positives, ensuring resources are directed to genuinely invaded areas. Recall (or sensitivity) is the proportion of true positives out of all actual presences (true positives plus false negatives). Focusing on recall emphasises capturing as much of the invasion range as possible. Depending on the species and available resources, stakeholders must balance precision and recall, prioritising precision to avoid unnecessary interventions or recall to ensure critical invaded areas are not overlooked (Sofaer et al. 2019). The F_β metric allows for adjustment of the relative importance placed on each, providing flexibility in stakeholder preference. Youden's J statistic (Youden 1950) is another strong op-

tion for use as an evaluation metric for selecting invasion boundary spatial scaling parameters (see Suppl. material 2 for an example analysis using Youden's J). This metric is particularly useful in this application because it is insensitive to class imbalances like the imbalance between presences and absences in the lydemapr database, which has more absences than presences. Using this metric, we find that slightly finer spatial resolutions were preferred than those selected using the F_β score with $\beta = 2$. For users uninterested in evaluating multiple β values or concerned about influencing results by making a subjective β value choice, the Youden's J statistic can be used instead.

The F_β metric originated in the field of information retrieval, where, for a given query, the number of irrelevant texts is exceedingly large (Christen et al. 2024). In such cases, it is valuable to focus on precision and recall, which are unaffected by the number of correctly classified negative cases. Similarly, there are innumerable locations where the target pest does not occur, and metrics that strongly reward correctly identified true negatives may misguide users working to find useful range limit boundaries. Instead, defining a range boundary that includes as many presences as possible (high recall), while excluding absences as much as possible (high precision) is desirable. Flexible weighting of these two goals is achieved by altering the β parameter in the F_β metric. We found that β values for the *L. delicatula* invasion of the US ranging from 1–3 generally produced clear peaks in optimal scaling parameters (cell edge length for grids and α for α -convex hulls, Fig. 3). Higher values of β , reflecting increased importance placed on capturing all presence records over limiting the inclusion of absences, resulted in coarser preferred α values. Very high β values did not indicate a clearly preferred scale, however, showing a broad range of α values that performed comparably, because once nearly all presences are being captured, including more absences has relatively little impact with high β values. Future work should examine how the choice of β might be constrained by data properties, such as class imbalance. Similarly, other metrics that balance precision and recall should also be explored and contrasted to F_β (Sofaer et al. 2019; Li and Guo 2021).

Our approach to locating *L. delicatula*'s invasion front was based on presence records made by state and federal agencies, and absences were used only to evaluate the boundaries drawn using these presences. This approach provides a highly detailed picture of the spread of this pest, but alternative methods may add new insights in the future. Uncertainty in the location of the invasion front is not constant over space: some locations are highly sampled, while others have only sparse sampling. For future efforts, hierarchical occupancy models may offer one tool to help quantify this uncertainty, though these approaches can be computationally demanding with large datasets like the one analysed here. Further, public reports of *L. delicatula* presence provide a wealth of information, but are likely to include some false positive reports that may hamper analysis, although methods for estimating the false positive rate are available (Clare et al. 2019). Our analysis also did not account for imperfect detection, where observers may record an absence when the targeted organism is truly present (Berec et al. 2015). Future work to estimate detection rates (Valentin et al. 2020) may help to more completely incorporate absence records into future analyses of *L. delicatula* spread dynamics.

We observed a slowdown in boundary displacement around the primary epicentre originating in Berks County, PA, in the years 2020 through 2023. The causes for this pattern remain unclear. The invasion reached the eastern shoreline

during this period. The invasion's margin also reached outside Pennsylvania for the first time during this period, and differences may, therefore, be caused in part by different sampling plans used by different state agencies (Fig. 1). The COVID-19 pandemic and shutdown also affected survey efforts during these years, reducing the sampling effort even as the invasion's border continued to grow (Table 1). The slowdown might also have been caused by biological factors. Climatic conditions might have been less favourable in these years than in 2017–2019 when boundary expansion was most rapid. As the invasion is relatively new, the sample size to make strong statements regarding climate's impact on spread rates, however, remains very low. Alternatively, accumulating natural enemies may slow population growth, thereby reducing spread (Johnson et al. 2023). It may also be true that more frequent establishment of small, isolated populations close to the margins of the primary invasion affected our measure of boundary displacement, which drew the shortest distance between one year's boundary and any occupied area in the prior year, not just to the targeted population's prior limit. With more small forerunner populations, drawn boundary displacement may be short, even if the main boundary moved a substantial distance.

Lycorma delicatula's spread in its invaded range in the USA is ongoing. Continuing establishment of far-flung outlying populations threatens to continue the rapid range expansion that we have observed in the years 2014 through 2023. Continued sampling effort is necessary to provide information for better mechanistic understanding of the ways in which this pest spreads. By looking at the history of spread, particularly the history of spread in the first few years after first detection, managers can derive useful projections for how far newly-established populations that might yet be targeted for eradication are likely to spread. Slow initial spread should motivate managers to take early action and treat nascent satellite populations before their spread rate accelerates. Additionally, analyses of invasive spread should carefully consider the scale used to derive estimates, as differences between scales can substantially alter estimated spread rates and the locations of invaded range boundaries. Effective surveying for *L. delicatula* and similar invasive species will require leveraging historical spread patterns, optimising spatial scales for analysis and incorporating cross-validation to improve range boundary accuracy. By combining these approaches with early intervention at nascent satellite populations, managers can mitigate the long-term impacts of invasive species.

Acknowledgements

For sharing data on *L. delicatula* occurrence, we thank the U.S. Department of Agriculture Spotted Lanternfly Working Group, the Pennsylvania Department of Agriculture, the New York State Department of Agriculture and Markets, the Massachusetts Department of Agricultural Resources and the North Carolina Dept. of Agriculture and Consumer Services. We thank Jocelyn Behm, Owen Schneider, Stefani Cannon, Jacob Woods, Robert Hendrickson, Lee Zimmerman and two anonymous reviewers for helpful feedback on this project.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Use of AI

No use of AI was reported.

Funding

This work was funded by the United States Department of Agriculture (USDA) Animal and Plant Health Inspection Service Plant Protection and Quarantine under agreements AP22PPQS&T00C146, AP22PPQS&T00C097, AP23PPQS&T00C090 and AP24PPQS&T00C136; the USDA National Institute of Food and Agriculture Specialty Crop Research Initiative under the coordinated agricultural project award 2019-51181-30014; the USDA Tactical Sciences for Agricultural Biosecurity under the project award 2022-68013-37139; the Pennsylvania Department of Agriculture under agreements C94000833 and C940001671.

Author contributions

All authors conceived the study and contributed to gathering data. JK carried out analyses and wrote the first draft of the manuscript. All authors contributed to revising the manuscript and approved the final draft.

Author ORCIDs

Joseph Keller  <https://orcid.org/0000-0001-7395-0582>

Sebastiano De Bona  <https://orcid.org/0000-0001-6193-7073>

Matthew R. Helmus  <https://orcid.org/0000-0003-3977-0507>

Data availability

Occurrence records of *Lycorma delicatula* are available in the lydemapr R data package (<https://github.com/ieco-lab/lydemapr>).

References

- Aikio S, Duncan RP, Hulme PE (2010) Lag-phases in alien plant invasions: separating the facts from the artefacts. *Oikos* 119: 370–378. <https://doi.org/10.1111/j.1600-0706.2009.17963.x>
- Allsopp PG (1996) Japanese beetle, *Popillia japonica* Newman (Coleoptera: Scarabaeidae): Rate of movement and potential distribution of an immigrant species. *The Coleopterists Bulletin* 50: 81–95.
- Andrew ME, Ustin SL (2010) The effects of temporally variable dispersal and landscape structure on invasive species spread. *Ecological Applications* 20: 593–608. <https://doi.org/10.1890/09-0034.1>
- Barringer LE, Ciafré CM (2020) Worldwide feeding host plants of spotted lanternfly, with significant additions from North America. *Environmental Entomology* 49: 999–1011. <https://doi.org/10.1093/ee/nvaa093>
- Belouard N, De Bona S, Helmus MR, Smith IG, Behm JE (2025) A method to quantify jump dispersal of invasive species from occurrence data: the case of the spotted lanternfly, *Lycorma delicatula*. *NeoBiota* 98: 319–334. <https://doi.org/10.3897/neobiota.98.147310>
- Berec L, Kean JM, Epanchin-Niell R, Liebhold AM, Haight RG (2015) Designing efficient surveys: spatial arrangement of sample points for detection of invasive species. *Biological Invasions* 17: 445–459. <https://doi.org/10.1007/s10530-014-0742-x>
- Burgman MA, Fox JC (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Animal Conservation* 6: 19–28. <https://doi.org/10.1017/S1367943003003044>

- Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach (2nd ed.). Springer, New York, 488 pp.
- Camac JS, Baumgartner JB, Hester S, Subasinghe R, Collins S (2021) Using edmaps & Zonation to inform multi-pest early-detection surveillance designs. Technical Report for CEBRA project 20121001.
- Chinchor N (1992) MUC-4 Evaluation Metrics. Proc. of the Fourth Message Understanding Conference, 22–29. <https://doi.org/10.3115/1072064.1072067>
- Christen P, Hand DJ, Kirielle N (2024) A review of the F-Measure: Its history, properties, criticism, and alternatives. ACM Computing Surveys 56: 1–24. <https://doi.org/10.1145/3606367>
- Clare JDJ, Townsend PA, Anhalt-Depies C, Locke C, Stenglein JL, Frett S, Martin KJ, Singh A, Van Deelen TR, Zuckerberg B (2019) Making inference with messy (citizen science) data: when are data accurate enough and how can they be improved? Ecological Applications 29: e01849. <https://doi.org/10.1002/eap.1849>
- Cook RT, Ward SF, Liebhold AM, Fei S (2021) Spatial dynamics of spotted lanternfly, *Lycorma delicatula*, invasion of the Northeastern United States. NeoBiota 70: 23–42. <https://doi.org/10.3897/neobiota.70.67950>
- Coutts SR, Helmstedt KJ, Bennett JR (2018) Invasion lags: The stories we tell ourselves and our inability to infer process from pattern. Diversity and Distributions 24: 244–251. <https://doi.org/10.1111/ddi.12669>
- Crall AW, Newman GJ, Stohlgren TJ, Holfelder KA, Graham J, Waller DM (2011) Assessing citizen science data quality: an invasive species case study. Conservation Letters 4: 433–442. <https://doi.org/10.1111/j.1755-263X.2011.00196.x>
- Crooks JA, Soulé ME (1999) Lag times in population explosions of invasive species: causes and implications. In: Sandlund OT, Schei PJ, Viken Å (Eds) Invasive Species and Biodiversity Management. Kluwer Academic Publishers, Dordrecht, Netherlands. https://doi.org/10.1007/978-94-011-4523-7_7
- De Bona S, Barringer L, Kurtz P, Losiewicz J, Parra GR, Helmus MR (2023) lydemapr: an R package to track the spread of the invasive spotted lanternfly (*Lycorma delicatula*, White 1845) (Hemiptera, Fulgoridae) in the United States. NeoBiota 86: 151–168. <https://doi.org/10.3897/neobiota.86.101471>
- Dewhurst S, Lutscher F (2009) Dispersal in heterogeneous habitats: thresholds, spatial scales, and approximate rates of spread. Ecology 90: 1338–1345. <https://doi.org/10.1890/08-0115.1>
- Edelsbrunner H, Kirkpatrick D, Seidel R (1983) On the shape of a set of points in the plane. IEEE Transactions on Information Theory 29: 551–559. <https://doi.org/10.1109/TIT.1983.1056714>
- Epanchin-Niell RS, Hastings A (2010) Controlling established invaders: integrating economics and spread dynamics to determine optimal management. Ecology Letters 13: 528–541. <https://doi.org/10.1111/j.1461-0248.2010.01440.x>
- Epanchin-Niell RS, Liebhold AM (2015) Benefits of invasion prevention: Effect of time lags, spread rates, and damage persistence. Ecological Economics 116: 146–153. <https://doi.org/10.1016/j.ecolecon.2015.04.014>
- Fahrner S, Aukema BH (2018) Correlates of spread rates for introduced insects. Global Ecology and Biogeography 27: 734–743. <https://doi.org/10.1111/geb.12737>
- Hastings A (1996) Models of spatial spread: A synthesis. Biological Conservation 78: 143–148. [https://doi.org/10.1016/0006-3207\(96\)00023-7](https://doi.org/10.1016/0006-3207(96)00023-7)
- Hastings A, Cuddington K, Davies KF, Dugaw CJ, Elmendorf S, Freestone A, Harrison S, Holland M, Lambrinos J, Malvadkar U, Melbourne BA, Moore K, Taylor C, Thomson D (2004) The spatial spread of invasions: new developments in theory and evidence: Spatial spread of invasions. Ecology Letters 8: 91–101. <https://doi.org/10.1111/j.1461-0248.2004.00687.x>
- Hauser CE, Giljohann KM, Rigby M, Herbert K, Curran I, Pascoe C, Williams NSG, Cousens RD, Moore JL (2016) Practicable methods for delimiting a plant invasion. Diversity and Distributions 22: 136–147. <https://doi.org/10.1111/ddi.12388>

- Hui C, Richardson DM, Robertson MP, Wilson JR, Yates CJ (2011) Macroecology meets invasion ecology: linking the native distributions of Australian acacias to invasiveness. *Diversity and Distributions* 17: 872–883. <https://doi.org/10.1111/j.1472-4642.2011.00804.x>
- Huron NA, Behm JE, Helmus MR (2022) Paninvasion severity assessment of a u.s. grape pest to disrupt the global wine market. *Communications Biology* 5: 1–11. <https://doi.org/10.1038/s42003-022-03580-w>
- IUCN Standards and Petitions Committee (2024) Guidelines for Using the IUCN Red List Categories and Criteria. Version 16. <https://www.iucnredlist.org/documents/RedListGuidelines.pdf>
- Johnson AE, Cornell A, Hermann S, Zhu F, Hoover K (2023) Using community science to identify predators of spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae), in North America. *Bulletin of Entomological Research* 113: 637–644. <https://doi.org/10.1017/S0007485323000317>
- Jones CM, Jones S, Petrasova A, Petras V, Gaydos D, Skrip MM, Takeuchi Y, Bigsby K, Meentemeyer RK (2021) Iteratively forecasting biological invasions with PoPS and a little help from our friends. *Frontiers in Ecology and the Environment* 19: 411–418. <https://doi.org/10.1002/fee.2357>
- Kreitman D, Keena MA, Nielsen AL, Hamilton G (2021) Effects of temperature on development and survival of nymphal *Lycorma delicatula* (Hemiptera: Fulgoridae). *Environmental Entomology* 50: 183–191. <https://doi.org/10.1093/ee/nvaa155>
- Ladin ZS, Eggen DA, Trammell TLE, D’Amico V (2023) Human-mediated dispersal drives the spread of the spotted lanternfly (*Lycorma delicatula*). *Scientific Reports* 13: e1098. <https://doi.org/10.1038/s41598-022-25989-3>
- Lampert A (2024) Optimizing strategies for slowing the spread of invasive species. *PLOS Computational Biology* 20: e1011996. <https://doi.org/10.1371/journal.pcbi.1011996>
- Lawson CR, Hodgson JA, Wilson RJ, Richards SA (2014) Prevalence, thresholds and the performance of presence–absence models. *Methods in Ecology and Evolution* 5: 54–64. <https://doi.org/10.1111/2041-210X.12123>
- Lewkiewicz SM, Seibold B, Helmus MR (2024) Quantifying population resistance to climatic variability: The invasive spotted lanternfly grape pest is buffered against temperature extremes in California. *Ecological Modelling* 497: 110841. <https://doi.org/10.1016/j.ecolmodel.2024.110841>
- Lewkiewicz SM, De Bona S, Helmus MR, Seibold B (2022) Temperature sensitivity of pest reproductive numbers in age-structured PDE models, with a focus on the invasive spotted lanternfly. *Journal of Mathematical Biology* 85: 1–29. <https://doi.org/10.1007/s00285-022-01800-9>
- Li W, Guo Q (2021) Plotting receiver operating characteristic and precision–recall curves from presence and background data. *Ecology and Evolution* 11: 10192–10206. <https://doi.org/10.1002/ece3.7826>
- Liang W, Tran L, Grant J, Srivastava V (2020) Estimating invasion dynamics with geopolitical unit-level records: The optimal method depends on irregularity and stochasticity of spread. *Sustainability* 12: 8526. <https://doi.org/10.3390/su12208526>
- Liang W, Tran L, Wiggins GJ, Grant JF, Stewart SD, Washington-Allen R (2019) Determining spread rate of kudzu bug (Hemiptera: Plataspidae) and its associations with environmental factors in a heterogeneous landscape. *Environmental Entomology* 48: 309–317. <https://doi.org/10.1093/ee/nvz014>
- Liebhold AM, Kean JM (2019) Eradication and containment of non-native forest insects: successes and failures. *Journal of Pest Science* 92: 83–91. <https://doi.org/10.1007/s10340-018-1056-z>
- Liebhold AM, Halverson JA, Elmes GA (1992) Gypsy moth invasion in North America: A quantitative analysis. *Journal of Biogeography* 19: e513. <https://doi.org/10.2307/2845770>
- Lonsdale WM (1993) Rates of spread of an invading species--mimosa pigra in Northern Australia. *The Journal of Ecology* 81: e513. <https://doi.org/10.2307/2261529>

- Memmott J, Craze PG, Harman HM, Syrett P, Fowler SV (2005) The effect of propagule size on the invasion of an alien insect. *Journal of Animal Ecology* 74: 50–62. <https://doi.org/10.1111/j.1365-2656.2004.00896.x>
- Meyer H, Reudenbach C, Wöllauer S, Nauss T (2019) Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. *Ecological Modelling* 411: e108815. <https://doi.org/10.1016/j.ecolmodel.2019.108815>
- Nathan R, Monk CT, Arlinghaus R, Adam T, Alós J, Assaf M, Baktoft H, Beardsworth CE, Bertram MG, Bijleveld AI, Brodin T, Brooks JL, Campos-Candela A, Cooke SJ, Gjelland KØ, Gupte PR, Harel R, Hellström G, Jeltsch F, Killen SS, Klefoth T, Langrock R, Lennox RJ, Lourie E, Madden JR, Orchan Y, Pauwels IS, Říha M, Roeleke M, Schlägel UE, Shohami D, Signer J, Toledo S, Vilc O, Westrelin S, Whiteside MA, Jarić I (2022) Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science* 375: eabg1780. <https://doi.org/10.1126/science.abg1780>
- Neubert MG, Parker IM (2004) Projecting rates of spread for invasive species. *Risk Analysis* 24: 817–831. <https://doi.org/10.1111/j.0272-4332.2004.00481.x>
- Pateiro-Lopez B, Rodriguez-Casal A (2022) alphahull: Generalization of the Convex Hull of a Sample of Points in the Plane.: 2.5. <https://doi.org/10.32614/CRAN.package.alphahull>
- Patton DR (1975) A diversity index for quantifying habitat “Edge.” *Wildlife Society Bulletin* (1973–2006) 3: 171–173.
- Pebesma E, Bivand R, Racine E, Sumner M, Cook I, Keitt T, Lovelace R, Wickham H, Ooms J, Müller K, Pedersen TL, Baston D, Dunnington D (2023) sf: Simple Features for R. [September 15, 2023] <https://doi.org/10.32614/cran.package.sf>
- R Core Team (2023) R: A Language and Environment for Statistical Computing. <https://doi.org/10.32614/r.manuals>
- Rejmanek M, Pitcairn MJ (2002) When is eradication of exotic pest plants a realistic goal? In: Veitch CR, Clout MN (Eds) *Turning the tide: The eradication of invasive species: proceedings of the International Conference on eradication of Island Invasives*. Occasional Paper of the IUCN Species survival Commission No. 27. IUCN, Gland, Switzerland, 249–253. <https://portals.iucn.org/library/efiles/documents/ssc-op-028.pdf>
- Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F, Dormann CF (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40: 913–929. <https://doi.org/10.1111/ecog.02881>
- Sharov AA, Roberts EA, Liebhold AM, Ravlin FW (1995) Gypsy Moth (Lepidoptera: Lymantriidae) Spread in the Central Appalachians: Three Methods for Species Boundary Estimation. *Environmental Entomology* 24: 1529–1538. <https://doi.org/10.1093/ee/24.6.1529>
- Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. *The American Naturalist* 146: 229–251. <https://doi.org/10.1086/285796>
- Shirey V, Belitz MW, Barve V, Guralnick R (2021) A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography* 44: 537–547. <https://doi.org/10.1111/ecog.05396>
- Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38: 196–218. <https://doi.org/10.1093/biomet/38.1-2.196>
- Sofaer HR, Hoeting JA, Jarnevich CS (2019) The area under the precision-recall curve as a performance metric for rare binary events. *Methods in Ecology and Evolution* 10: 565–577. <https://doi.org/10.1111/2041-210X.13140>
- Storer TI (1937) The muskrat as native and alien. *Journal of Mammalogy* 18: e443. <https://doi.org/10.2307/1374334>

- Tisseuil C, Gryspeirt A, Lancelot R, Pioz M, Liebhold A, Gilbert M (2016) Evaluating methods to quantify spatial variation in the velocity of biological invasions. *Ecography* 39: 409–418. <https://doi.org/10.1111/ecog.01393>
- Tobin PC, Robinet C (2022) Advances in understanding and predicting the spread of invading insect populations. *Current Opinion in Insect Science* 54: e100985. <https://doi.org/10.1016/j.cois.2022.100985>
- Tobin PC, Liebhold AM, Anderson Roberts E (2007) Comparison of methods for estimating the spread of a non-indigenous species. *Journal of Biogeography* 34: 305–312. <https://doi.org/10.1111/j.1365-2699.2006.01600.x>
- Tobin PC, Blackburn LM, Gray RH, Lettau CT, Liebhold AM, Raffa KF (2013) Using delimiting surveys to characterize the spatiotemporal dynamics facilitates the management of an invasive non-native insect. *Population Ecology* 55: 545–555. <https://doi.org/10.1007/s10144-013-0382-5>
- Triska MD, Renton M (2018) Do an invasive organism’s dispersal characteristics affect how we should search for it? *Royal Society Open Science* 5: 171784. <https://doi.org/10.1098/rsos.171784>
- Urban JM, Leach H (2023) Biology and Management of the Spotted Lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae), in the United States. *Annual Review of Entomology* 68: 151–167. <https://doi.org/10.1146/annurev-ento-120220-111140>
- Urban JM, Calvin D, Hills-Stevenson J (2021) Early response (2018–2020) to the threat of spotted lanternfly, *Lycorma delicatula* (Hemiptera: Fulgoridae) in Pennsylvania. *Annals of the Entomological Society of America* 114: 709–718. <https://doi.org/10.1093/aesa/saab030>
- Urban MC, Phillips BL, Skelly DK, Shine R (2008) A toad more traveled: The heterogeneous invasion dynamics of cane toads in Australia. *The American Naturalist* 171: E134–E148. <https://doi.org/10.1086/527494>
- Valentin RE, Fonseca DM, Gable S, Kyle KE, Hamilton GC, Nielsen AL, Lockwood JL (2020) Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. *Molecular Ecology Resources* 20. <https://doi.org/10.1111/1755-0998.13151>
- Wallace RD, Barger CT, Reaser JK (2020) Enabling decisions that make a difference: guidance for improving access to and analysis of invasive species information. *Biological Invasions* 22: 37–45. <https://doi.org/10.1007/s10530-019-02142-2>
- Ward SF, Fei S, Liebhold AM (2020) Temporal dynamics and drivers of landscape-level spread by emerald ash borer. *Journal of Applied Ecology* 57: 1020–1030. <https://doi.org/10.1111/1365-2664.13613>
- Youden WJ (1950) Index for rating diagnostic tests. *Cancer* 3: 32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::aid-cnrcr2820030106>3.0.co;2-3;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cnrcr2820030106>3.0.co;2-3;2-3)

Supplementary material 1

Supplementary tables and figures

Authors: Joseph Keller, Sebastiano De Bona, Matthew R. Helmus

Data type: docx

Explanation note: **table S1.1:** Regression results, with invaded area determined based on intersection of established points with grids of various scales, fit using simple linear regression. **table S1.2:** Regression results, with invaded area determined based on intersection of established points with grids of various scales, fit using nonlinear, logistic regression. **table S1.3:** Regression results, with invaded area determined based on the α -convex hull around established points, fit using simple linear regression. **table S1.4:** Regression results, with invaded area determined based on the α -convex hull around established points, fit using nonlinear, logistic regression. **table S1.5:** Estimated radial spread rate (km/yr) of *L. delicatula* based on effective range radius regression with the invaded range determined by occupied grid cells of varying scale. **table S1.6:** Estimated radial spread rate (km/yr) of *L. delicatula* based on effective range radius regression with the invaded range determined by α -convex hulls of varying scale. **fig. S1.1:** Cross validation allows for the selection of a best performing spatial scale for analysis. **fig. S1.2:** Spread rate varied with satellite population age.

Copyright notice: This dataset is made available under the Open Database License (<http://opendata-commons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.106.177041.suppl1>

Supplementary material 2

Supplementary figures

Authors: Joseph Keller, Sebastiano De Bona, Matthew R. Helmus

Data type: docx

Explanation note: **fig. S2.1:** The spatial scaling parameter α' affected the false positive rate and sensitivity of boundaries drawn around the *L. delicatula* invasion in each year of its progression 2015–2023. **fig. S2.2:** The grid cell edge length affected the false positive rate and sensitivity of boundaries drawn around the *L. delicatula* invasion in each year of its progression 2015–2023. **fig. S2.3:** The best-performing α -hull outperformed the best-performing grid in each year of the *L. delicatula* invasion 2015–2023 as measured using Youden's J. Differences in the Youden's J values were calculated by subtracting the calculated value for the value for the best performing grid from the best performing α -hull in each year.

Copyright notice: This dataset is made available under the Open Database License (<http://opendata-commons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/neobiota.106.177041.suppl2>